Orthogonal Chaotic Vector Shift Keying in Digital Communications

Timothy J. Wren and Tai C. Yang,
Department of Engineering and Design, University of Sussex, Brighton, BN1 9QT, UK.

Abstract —

An Orthogonal Chaotic Vector Shift Keying (OCVSK) digital communication scheme is presented in this paper. The main characteristics of the scheme are increased data transmission rates with greatly improved robustness and, an increase in security of communications links due to the structure of the scheme and the nature of the message bearer. Compared with some existing schemes, for example Quadrature Chaos Shift Keying (QCSK) reported in the literature, the noise rejection is improved by an increase in the 'inter-symbolic separation'. Furthermore, a new method of characterizing non-linear processing elements in complex communication schemes has been presented. Based on this, a simple modelling and evaluation method to determine the Bit Error Rates (BER) of these schemes is derived. Various simulated results are presented to demonstrate these achievements.

I. INTRODUCTION

Using chaotic signals for secure communications has attracted some considerable research interests, for example see [1-7] and references therein. Chaotic systems are stable but oscillate aperiodically around a strange attractor, making them naturally harder to identify and to predict. Chaotic communication signals are spread spectrum in nature which utilizes a large bandwidth and has a low power spectral density. In traditional communication systems, the analogue sample functions sent through the channel are weighted sums of sinusoidal waveforms and are linear. However, in chaotic communication systems, the samples are segments of chaotic waveforms and are nonlinear. This nonlinear stable aperiodic characteristic of chaotic communication has numerous features that make it attractive for communication use. It has a wideband characteristic, it is resistant against multipath fading and it offers a cheaper solution to traditional spread spectrum systems. In chaotic communications, the digital information to be transmitted is placed directly onto a wide-band chaotic signal. Among several systems proposed, one of the best Bit Error Rate (BER) performances has been achieved by the Differential Chaos Shift Keying (DCSK) scheme and its variations [17, 18, and 19]. These schemes are based upon wideband chaotic signals which, under severe multipath propagation, exhibit a better performance than conventional systems based on sinusoidal carriers.

DCSK is a transmitted-reference digital signaling scheme [20]. For each symbol period, the DCSK signal consists of a piece of chaotic waveform, followed by its non-inverted or inverted copy, depending on the binary symbol ("0" or "1") to be transmitted. The first and the second part of the DCSK signal are called the reference and the information-bearing chip, respectively. Several different methods have been proposed in the literature to increase the data rate of DCSK [21, 22]. The simplest option consists of scaling the information and/or the reference parts of the signal. For example, the information bearing part may be multiplied by a number depending on the symbol transmitted. A more sophisticated approach uses two chaotic basis functions and divides the symbol period into four time slots in order to obtain a multilevel scheme [21]. These methods, though, achieve higher data rate by giving up some of the BER performance. A simple idea of transmitting more than one information bearing chip for one reference chip in order to improve the performance of DCSK was proposed in [23]. A novel multilevel chaos-based communication scheme is proposed in [7], called Quadrature Chaos Shift Keying (QCSK). It is characterized by the same bandwidth occupation and similar BER performance as DCSK, but has a higher data rate. QCSK may be considered as the chaotic counterpart of quadrature phase shift keying (QPSK) in conventional digital communications. QPSK exhibits the same BER performance as binary phase shift keying (BPSK) with the same bandwidth occupation, but double data rate. This is achieved by employing

a quadrature pair of sinusoidal carriers to generate an orthogonal signal basis. Since the basis components are orthogonal, they can be used to modulate information separately as for two BPSK systems sharing the same channel without interfering with each other. Orthogonal basis functions, usually sinusoids, are used in digital communications to generate large signal constellations in order to increase the spectral efficiency. Typical examples are *M*-ary Phase Shift Keying (PSK) where the phase of the transmitted signal is varied among *M* discrete values and quadrature amplitude modulation (QAM) where both the amplitude and the phase of the reference sinusoid are varied [24]. The basic idea underlying the QCSK scheme is the generation of chaotic signals which are orthogonal over a specified time interval. This allows the creation of a basis of chaotic functions from which arbitrary constellations of chaotic signals can be constructed. For instance, in QCSK, a linear combination of two chaotic basis functions is used to encode four symbols. The key point for exploiting this idea in a communication system is that one must be able to generate the chaotic basis functions starting from a single chaotic signal. The same concept holds for conventional digital communication schemes such as QPSK, where the quadrature component can be obtained from the in phase one by means of a simple phase shifter.

In this paper, with the similar underlying ideas of QCSK, an Orthogonal Chaotic Vector Shift Keying (OCVSK) scheme is proposed.

The basis functions for OCVSK are linear combinations of orthogonal pieces of chaotic waveforms derived from a reference sequence by using the Gram Schmidt method on sampled signal vector sequences. It is shown that it is a non-trivial problem to extend the Fourier analysis and Hilbert Transform approach adopted by the QCSK method to higher dimensions and that the historically based Fourier analysis representations are not a true reflection of the nature of the employed signals. The complex representation is derived from complex analysis, which employs the inherent orthogonality of the sine and cosine functional representation. If symbols are represented within a higher dimensional space, then the apparent dependence on inherent orthogonality can be discarded; the effective separation between the symbols can be increased and the effects of noise reduced. With the feasibility of separating the problem from complex analysis the problem can be reduced to finding sets of mutually orthogonal signals.

The OCVSK method has significant improvements in transmission efficiency over some existing methods. The information transmitted per unit time is shown to be dependent on the dimensions adopted under the new method. Similarly, the noise rejection is greatly improved due to the increased “inter-symbolic separation”. Therefore, the new method tends to be more robust. Within the physical limits of communication channels, the new method provides a way of increasing the security of digital communications. Also presented is a novel generic method for characterizing and simply modelling the noise transmission behaviours of communication schemes, including the OCVSK method proposed in this paper. In addition, an analytical formula of the Bit Error Rates for any scheme has been derived.

Section II provides the necessary background materials [7] related to the proposed OCVSK scheme. It discusses the limitations of the two dimensional schemes considered in [7]. The theoretical part of the new scheme is presented in section III. A particular architecture of the scheme, with encoding and decoding methods, is described in section IV. This is followed, in section V, by a quantitative analysis of the noise on the transmitted signals within the proposed scheme. Furthermore, a generic method to calculate the signal to noise ratio is derived in sections VI and is then applied to the proposed scheme. Section VII presents a case study in which a dimensional value of four has been chosen; this value is sufficiently high to demonstrate the advantages of the scheme whilst presenting a clear set of results. Finally, section VIII concludes the paper.

II. BACKGROUND

The OCVSK scheme is based on a combination of the Differential Chaos Shift Keying (DCSK) method described in [8], and a derivation of the well-known Quadrature Phase Shift Keying (QPSK), which itself is the quadrature form of Binary Phase Shift Keying (BPSK) [9]. In both of the latter, the underlying message bearer is sinusoidal. For the BPSK technique a portion of the sinusoidal signal is transmitted to represent a ‘0’ and its anti-phase counterpart is transmitted to represent a ‘1’. QPSK requires two orthogonal signals, which are added together in a combination of four ways to give a four state transmitted

signal. The term orthogonal in this sense means, that the integral over a fixed period of the product of two functions, has a mean of zero. That is

$$\frac{1}{T}\int_0^T x(t)y(t)dt = 0 \tag{1}$$

In the receiver, the signal parameters are determined by correlating the received signal with each of the orthogonal reference signals, and hence the exact meaning of the received signal can be interpreted. The signals used in this technique are sinusoidal and their orthogonal counterparts are cosine functions.

A. Quadrature Chaos Shift Keying

For differential methods, such as DCSK, each symbol is transmitted in two parts. The first element is a reference signal and the second is the message bearer. In DCSK the reference is a chaotic signal, generated by some chaotic process, and the message element is a replica of this to represent a '0' or an anti phase element representing a '1'.

In QCSK the sinusoidal signals are again replaced by chaotic reference signals. Signals that are orthogonal to them are then generated, and these signals are used in a similar way to the QPSK set of orthogonal signals. An example of a set of two orthogonal signals is shown in figure 1 where (a) is the chaotic signal and (b) is its orthogonal counterpart. Appendix A outlines the theory behind the QCSK scheme.

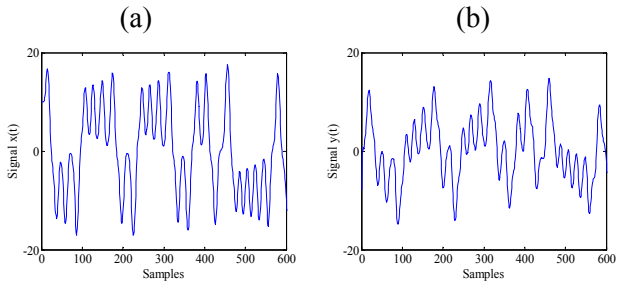

Figure 1:
QCSK Orthogonal Signal Set (a) Original Chaotic Signal and (b) the Orthogonal Chaotic Counterpart

Consider now two possible maximally separated equal amplitude constellations of signals that consist of an addition of a proportion of each orthogonal signal. These can be represented on an Argand diagram shown in figure 2 and the encoding values set out in table 1

	Symbol	0	1	2	3
(a)	**c**	1 0	0 1	-1 0	0 -1
(b)	**c**	$1/\sqrt{2}$ $1/\sqrt{2}$	$-1/\sqrt{2}$ $1/\sqrt{2}$	$-1/\sqrt{2}$ $-1/\sqrt{2}$	$1/\sqrt{2}$ $-1/\sqrt{2}$

Table 1

Each symbol can be represented by a complex number as

$$c = c_r + jc_i \tag{2}$$

Orthogonality is assured in the complex plane and its analogy can therefore be represented along the real time axis, if the two complimentary signals are orthogonal; that is

$$s(t) = c_r x(t) + c_i y(t) \tag{3}$$

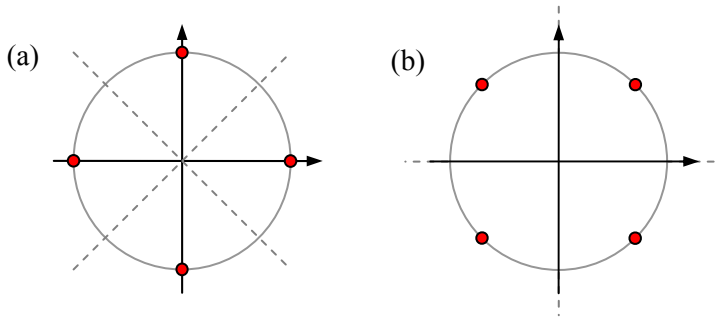

Figure 2: Maximal Separation Quadrature Constellations
(a) Symbol encoding contains the reference signal whereas, (b) Encodes all symbols.

this is the message signal for each symbol in the message. At the receiver the symbols can be retrieved by determining the coefficients of each individual orthogonal component by using the two correlation integrals

$$c_r = \frac{1}{P_x T}\int_0^T s(t)x(t)dt \tag{4}$$

$$c_i = \frac{1}{P_y T}\int_0^T s(t)y(t)dt \tag{5}$$

B. Limitations of Two Dimensional Schemes with Increased Transmission Efficiency

Here the limitations of Quadrature Chaos Shift Keying (QCSK) type methods are explored. These methods have been developed into an M-ary constellation methods, allowing the transmission of more symbols, thus improving the symbol transmission efficiency. This is illustrated below in figure 3 (a) and (b).

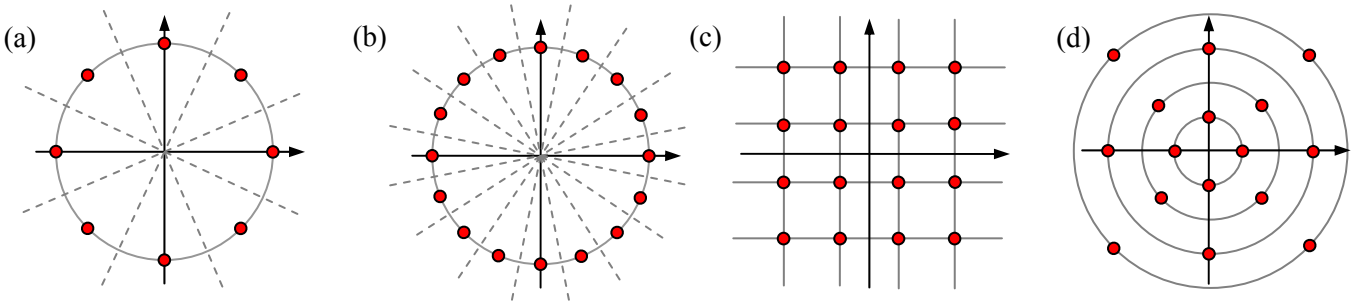

Figure 3: M-ary Constellations (a) ‘8’ Symbol ‘3’ bit and (b) ‘16’ Symbol ‘4’ bit representations.
QAM Constellations (c) ‘16’ Symbol ‘4’ Bit Grid and (d) ‘16’ Symbol ‘4’ Bit Circular Grid

The principal disadvantage with the first method is that all of the points in the constellation lie on a fixed radius circle represented on the complex plane. Large numbers of symbols require an equally large number of points on the fixed circle, which becomes crowded, and consequently gives rise to potential misinterpretation on decoding; these extended schemes become progressively less robust with an increase in bit number representation. One way to avoid this is to vary both the amplitude of the symbol representations as well as the phase. This is the well known Quadrature Amplitude Modulation (QAM) illustrated in figure 3(c) and (d).

This form of variation of grid type or circle radius type constellation gives rise to signal amplitude variation, which is generally not desirable from a security point of view as the signals are more easily detectable due to the varying power of the transmitted signals. QCSK, QAM and other types of communication constellation scheme are usually represented on the complex plane. This is largely

historical and based on Fourier analysis representation and not a true reflection of the nature of the employed signals. The complex representation is derived from complex analysis, which employs the inherent orthogonality of the sine and cosine functional representation. If symbols are represented within a higher dimensional space, then the apparent dependence on inherent orthogonality can be discarded; the effective separation between the symbols can be increased and the effects of noise reduced. Appendix B demonstrates that it is non-trivial problem to find a set of mutually orthogonal signals derived by Fourier analysis. It can therefore be concluded that the Fourier expansion method of orthogonal signal generation is not applicable to dimensions greater than m=2. So the problem can be restated as finding sets of mutually orthogonal signals. The combination of these signals can, in a similar way to QCSK, be extended to a much greater information capacity and hence transmission efficiency. QCSK introduces this idea, but is immediately constrained, by the use of the Fourier expansion and Hilbert Transform methods.

III. THEORY

Consider now a system, with an m dimensional constellation, that relies on m different mutually orthogonal signals, which actually form part of an orthonormal basis of functions $u_i(t)\ \forall\ i \in [1,m]$. As with the QCSK scheme, the message can be encoded using these orthogonal functions by combining them linearly using the value of the encoding coefficients. This can be represented as

$$s(t) = c_1 u_1(t) + c_2 u_2(t) + c_3 u_3(t) + \ldots + c_m u_m(t) \tag{6}$$

this in vector function notation becomes

$$s(t) = \mathbf{u}^T(t)\mathbf{c} \tag{7}$$

where

$$\mathbf{u}^T(t) = [u_1(t), u_2(t), u_3(t), \ldots, u_m(t)] \tag{8}$$

and

$$\mathbf{c}^T = [c_1, c_2, \ldots, c_m] \tag{9}$$

this is the message signal for each symbol in the transmitted message.

At the receiver the symbols can be retrieved by determining the coefficients of individual orthogonal components by using the m correlation integrals

$$c_i = \frac{1}{P_i T}\int_0^T s(t)u_i(t)dt \tag{10}$$

$$P_i = \frac{1}{T}\int_0^T u_i^2(t)dt \quad \forall \quad i \in [1,m] \tag{11}$$

or from equations 7-11

$$\int_0^T \mathbf{u}(t)s(t)dt = \int_0^T \mathbf{u}(t)\mathbf{u}^T(t)\mathbf{c}\,dt \tag{12}$$

Therefore this can be written in vector notation as

$$\mathbf{c} = \left[\int_0^T \mathbf{u}(t)\mathbf{u}^T(t)dt\right]^{-1}\int_0^T \mathbf{u}(t)s(t)dt \tag{13}$$

This will work with any set of signals if they are independent. If there is no noise present, the signal sets are orthogonal and the inversion is simplified by the inverted matrix being diagonal. However in the presence of noise, the inversion can influence the noise rejection characteristics of decoding. Noise rejection can be improved by discarding non diagonal terms because they are perceived to have been

generated by noise. The scheme of signal transmission here is irrelevant to the above derivation. The signals can be transmitted simultaneously on multiple channels or contiguously on one channel.

A. Generation of Orthogonal Signal Sets

The generation of a set of *m* orthogonal signals is required; we can approach this problem by first considering an *n* dimensional space. Any point $\mathbf{p}$ can be represented by an *n* dimensional vector that is a linear sum of the set of orthonormal basis vectors $\mathbf{u}_i$ where $\mathbf{u}_i \in R^n$ and $i \in [1, n]$.
Therefore the following can be written

$$\mathbf{p} = p_1\mathbf{u}_1 + p_2\mathbf{u}_2 + p_3\mathbf{u}_3 + ... + p_n\mathbf{u}_n \tag{14}$$

where $p_i \ \forall \ i \in [1, n]$ are real coefficients.
Now consider a subset of size *m* of these basis vectors that describe an *m* dimensional subspace within the *n* dimensional space. Further consider the set of vectors describing some hypersurface **s** within this *m* dimensional subspace.

$$\mathbf{s} = c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + c_3\mathbf{u}_3 + ... + c_m\mathbf{u}_m \tag{15}$$

Here $\mathbf{p}$ and $\mathbf{s} \in R^n$ and $m \le n$. The size of *n* is explored in section VII but is typically an order of magnitude greater in size than *m* for noise rejection purposes. This can be seen as analogous to equation 6, except that the summation is in terms of real vectors, and not real functions of *t*. The real vectors $\mathbf{u}_i$ can be obtained from real orthogonal functions $u_i(t)$ by a transformation described in appendix C producing a matrix $\mathbf{U} = [\mathbf{u}_1 \cdots \mathbf{u}_m]$ from a sampled matrix $\mathbf{X} = [\mathbf{x}_1 \cdots \mathbf{x}_m]$ by a simple linear transformation by an upper triangular matrix **W**. That is

$$\mathbf{X} = \mathbf{UW} \tag{16}$$

where

$$\mathbf{U}^T\mathbf{U} = \mathbf{I}_m \tag{17}$$

and **W** can be found as

$$\mathbf{U}^T\mathbf{UW} = \mathbf{U}^T\mathbf{X} \tag{18}$$

So given equations 16-18 it follows that

$$\mathbf{W} = \mathbf{U}^T\mathbf{X} \tag{19}$$

However, finding this transformation is unnecessary if the Gram-Schmidt algorithm is used; it is specified here because it will be needed for signal characterization in section VI.

For the proposed scheme to work each signal sequence needs to be independent of the last $m-1$ other signal sequences. So consider the independence of the columns of $\mathbf{X}$ made up from the last *m* sequences, and how this relates to the potential number of bits of precision that the signal set values may be in error. In order to determine if the last signal sequence is 'good' enough for transmission, an estimate of the 'bits' in precision error (B_e) can be determined from the matrix 2-norm condition number C_n as

$$B_e \propto \log_2(C_n) \tag{20}$$

where the condition number is given as

$$C_n = \max\left(\sqrt{\lambda}_i\right) / \min\left(\sqrt{\lambda}_i\right) \quad \forall \quad i \in [i, m] \tag{21}$$

here the λ_i represent the eigenvalues of the symmetric matrix $\mathbf{X}^T\mathbf{X}$ and $\sqrt{\lambda_i}$ are the singular values of **X.** If B_e is larger than the bit precision of the signal set values that are to be transmitted the signal sequence can be rejected. Simulation has shown that the probability of sequence rejection is very low.

IV. SYSTEM ARCHITECTURE: ENCODING AND DECODING SCHEMES

There are a number of different architectures which could be employed to exploit this communication scheme. Presented is a scheme called the 'Indirect Persistent **x** Scheme' because it is directly comparable to the DCSK and QCSK schemes outlined in section II. This scheme, shown in figure 4, has a greatly increased transmission efficiency and improved noise rejection dependent on the chosen dimension.

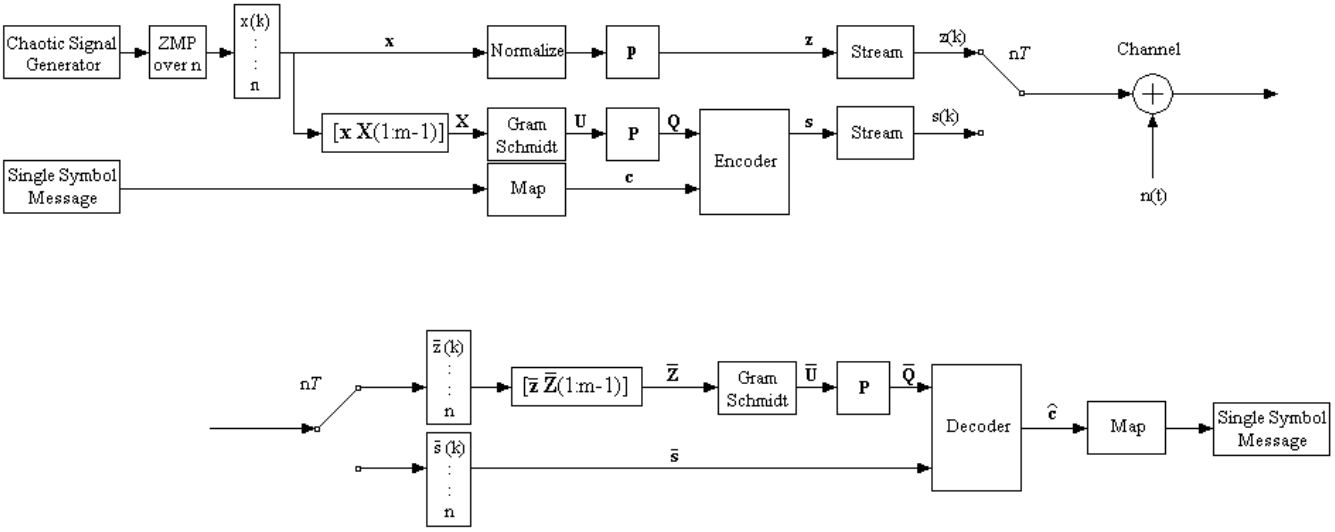

Figure 4: Indirect Persistent '**x**' Scheme System Architecture showing the Encoding and Decoding Scheme

A. Indirect Persistent 'x' Scheme

Consider the n length signal vector **x** produced by taking n samples of a chaotic process. Each sampled vector has the mean value removed thus leaving it as samples of a zero mean process. An $n \times m$ matrix **X** is formed from collections of n length **x** vectors. Each **x** vector of these collections remains persistent within the encoding architecture over m symbolic transmissions. Each symbol sequence transmits m bits of information so the transmission efficiency of this scheme is high, because the symbolic and bit data rate is only dependent on the length of each signal vector. Now an $n \times m$ orthonormal matrix **U** is generated from the **X** matrix using the Gram-Schmidt process. The matrix **U** is multiplied by a diagonal power balancing matrix **P** to produce a matrix **Q** as

$$\mathbf{Q} = \mathbf{UP} \tag{22}$$

The choice of the diagonal matrix **P** is arbitrary but specifying it in terms of a signal to noise power ratio will become significant in section VII. The same diagonal value of the **P** matrix is used to power balance the most recent normalized **x** vector before it is transmitted as the **z** vector. That is

$$\mathbf{z} = \tilde{\mathbf{x}}p \tag{23}$$

where $\tilde{\mathbf{x}}$ here is the normalized form of **x** that is

$$\tilde{\mathbf{x}}^T \tilde{\mathbf{x}} = 1 \tag{24}$$

A new **X** matrix is created after each **x** vector is sampled as

$$\mathbf{X}_{n,m} = [\mathbf{x} \quad \mathbf{X}_{n,m-1}] \tag{25}$$

It is now necessary only to encode a single symbol after each n samples represented by an **s** vector as

$$\mathbf{s} = \mathbf{Qc} \tag{26}$$

and this is transmitted in that same way as the **z** vector.

A brief description of a generalized method of selecting the **c** vector is now outlined. The **c** vector is a real valued vector of length *m* and represents the proportion of each of the orthogonal signal sequences that makes up the message bearing sequence. In the simplest form, which is comparable to QCSK, each of the components have a positive or negative value that makes **c** a unit vector. The potential set of encoding vectors $\mathbf{c}_j \ \forall \ j \in [1,2^m]$ is chosen from an encoding map $\mathbf{C} \in R^{m \times 2^m}$ which is calculated as

$$\mathbf{C} = [\mathbf{c}_1, \mathbf{c}_2, ..., \mathbf{c}_{(2^m)}] \tag{27}$$

$$\mathbf{c}_j = \frac{(2 \cdot \mathbf{b}(j-1) - \boldsymbol{\mu}_m)}{\sqrt{m}} \tag{28}$$

where the $\mathbf{c}_j$ vectors lie on an *m* dimensional unit hypersphere within the *n* space; $\boldsymbol{\mu}_m \in R^m$ is a vector $\mu_i = 1 \ \forall \ i \in [1,m]$ and $\mathbf{b}(j-1) \ \forall \ j \in [1,2^m]$ is the bit pattern function converting a bit pattern to an ordered vector, that is

$$\mathbf{b}(j) = [b_1 \ b_2 \cdots b_m]^T \tag{29}$$

and

$$j = 1 + \sum_{i=1}^{m} b_i 2^{m-i} \qquad b_i \in [0,1] \ \forall \ i \in [1,m] \tag{30}$$

Consider now the method for decoding each received signal vector, which is the equivalent of the correlation integral in equation 13, and is a least squares approximation of the encoding vector given a noisy received signal vector $\bar{\mathbf{s}}$.

$$\hat{\mathbf{c}} = \left[\overline{\mathbf{Q}}^T \overline{\mathbf{Q}}\right]^{-1} \overline{\mathbf{Q}}^T \bar{\mathbf{s}} \tag{31}$$

The derivation of equation 31 is given in appendix D.

The $\overline{\mathbf{Q}}$ matrix needs to be estimated from the persistent received reference matrix $\overline{\mathbf{Z}}$ created as

$$\overline{\mathbf{Z}}_{n,m} = [\bar{\mathbf{z}} \quad \overline{\mathbf{Z}}_{n,m-1}] \tag{32}$$

now both $\overline{\mathbf{U}}$ and $\overline{\mathbf{Q}}$ matrices can be formed by using the $\overline{\mathbf{Z}}$ matrix via the Gram-Schmidt process. The decoding equation 31 can now be simplified by substituting the received form of equation 22 into equation 31 to give

$$\hat{\mathbf{c}} = \mathbf{P}^{-1} \overline{\mathbf{U}}^T \bar{\mathbf{s}} \tag{33}$$

where

$$\mathbf{P}^{-1} = \frac{1}{p} \mathbf{I}_m \tag{34}$$

and p is the power balancing gain.

This scheme has a robust estimating structure because it avoids the noise transmission through an *m* dimensional matrix inversion and it has the same dependency on the nature of the noise transmission through the Gram-Schmidt process. The cyclic transmission efficiency is increased and is scalable with the dimension *m,* without any noise or time penalties.

V. SIGNAL CHARACTERIZATION

In this section, a generic characterization method is presented which considers the effect of noise transmitted through the various processes in the estimators.

A. BER Probability Formulation

To find the Bit Error Rate (BER), as a function of the number of samples *n* for each bit and the signal power to noise power ratio P_{snr} of the system, the following probability formulation will enable a simple method of BER calculation to be developed. The BER is considered as the probability that a singular

transmitted bit is decoded incorrectly in the receiver. This can be considered as a function of the probability of the estimate of the encoding vector lying outside its permitted region.
Consider then the probability of getting all bits correct that is

$$P(C_\mu) = 1 - P(E_\mu) \tag{35}$$

where $P(E_\mu)$ is the probability of any error and μ is the number of symbols, hence if b is the number of bits representing μ symbols then

$$\mu = 2^b \tag{36}$$

The probability of the i^{th} symbol being correct is

$$P(C_i) = \sqrt[b]{P(C_\mu)} \tag{37}$$

this gives the probability of a symbol error as

$$P(E_i) = 1 - P(C_i) \tag{38}$$

Now this is equivalent to the Bit Error Rate so from equations 35 to 38

$$BER = 1 - \sqrt[b]{1 - P(E_\mu)} \tag{39}$$

*B. Indirect m Symbol '**x**' Scheme Characterization*

The transmittable signal matrix $\mathbf{Z}$ is received in a noise contaminated form as $\overline{\mathbf{Z}}$. A noise contaminated orthonormal set of signal vectors $\overline{\mathbf{U}}$ can be derived using the Gram-Schmidt process and an equivalent $\overline{\mathbf{Q}}$ matrix can be found as

$$\overline{\mathbf{Q}} = \overline{\mathbf{U}}\mathbf{P} \tag{40}$$

The $\overline{\mathbf{U}}$ matrix has properties that can be characterized, by considering it to be expressed as a series of column vectors with the same power. As $\overline{\mathbf{U}}$ consists of a partial orthonormal basis over a limited m dimensional span of the n space it can be characterized as

$$\overline{\mathbf{U}} = \left[\frac{\mathbf{I}_m}{\mathbf{0}_{n-m,m}} \right] \tag{41}$$

This signal aggregation is valid, but does not describe the errors induced by the presence of noise on the Gram-Schmidt process. Equation 41 is clearly independent of any process and is notionally derived in this form to ensure it has the correct properties principally that

$$\overline{\mathbf{U}}^T \overline{\mathbf{U}} = \mathbf{I}_m \tag{42}$$

Appendix E derives the following characterization equations

$$\overline{\mathbf{Q}} = \sigma\sqrt{nP_{snr}} \cdot \mathbf{G}\left(\sqrt{nP_{snr}} \cdot \mathbf{UW} + \mathbf{E}\right) \tag{43}$$

$$\overline{\mathbf{s}} = \sigma\left(\sqrt{nP_{snr}} \cdot \left[\frac{\mathbf{I}_m}{\mathbf{0}_{n-m,m}} \right]\mathbf{c} + \boldsymbol{\varepsilon}\right) \tag{44}$$

VI. SIGNAL TO NOISE CALCULATIONS

In order to evaluate the performance of the transmission scheme against other schemes, a novel method is presented for producing Bit Error Rate results using the probability formulation and the signal characterization of section V. The results of these formulations are presented in section VII.

*A. Indirect Persistent '**x**' Scheme*

The symbol encoding vector estimate can be expressed from equation 31

$$\hat{\mathbf{c}} = [\overline{\mathbf{Q}}^T \overline{\mathbf{Q}}]^{-1} \overline{\mathbf{Q}}^T \overline{\mathbf{s}}$$

this can be expressed in terms of the noise on the transmission channel and the original idealised symbol encoding vector by the substitution of equations 43 and 44, that is

$$\begin{aligned}\overline{\mathbf{Q}} &= \sigma\sqrt{nP_{snr}}\cdot\mathbf{G}\left(\sqrt{nP_{snr}}\cdot\left[\frac{\mathbf{I}_m}{\mathbf{0}_{n-m,m}}\right]\mathbf{W}+\mathbf{E}\right)\\ &= \sigma\sqrt{nP_{snr}}\cdot\overline{\mathbf{U}}\end{aligned} \tag{45}$$

this yields

$$\begin{aligned}\hat{\mathbf{c}} &= \left(\left(\sigma\sqrt{nP_{snr}}\cdot\overline{\mathbf{U}}\right)^T\left(\sigma\sqrt{nP_{snr}}\cdot\overline{\mathbf{U}}\right)\right)^{-1}\cdot\sigma\left(\sqrt{nP_{snr}}\cdot\overline{\mathbf{U}}\right)^T\sigma\left(\sqrt{nP_{snr}}\cdot\left[\frac{\mathbf{I}_m}{\mathbf{0}_{n-m,m}}\right]\mathbf{c}+\boldsymbol{\varepsilon}\right)\\ &= \overline{\mathbf{U}}^T\left(\left[\frac{\mathbf{I}_m}{\mathbf{0}_{n-m,m}}\right]\mathbf{c}+\frac{1}{\sqrt{nP_{snr}}}\boldsymbol{\varepsilon}\right)\end{aligned} \tag{46}$$

where

$$\overline{\mathbf{U}} = \mathbf{G}\left(\sqrt{nP_{snr}}\cdot\left[\frac{\mathbf{I}_m}{\mathbf{0}_{n-m,m}}\right]\mathbf{W}+\mathbf{E}\right) \tag{47}$$

The result here is constructed using the power of the signal to noise ratio P_{snr}, whereas most of the literature quotes the equations and results in terms of the 'energy per bit divided by the noise power', that is $\frac{E_b}{N_0}$. This depends on the bit transmission rate, which in turn, is dependent on the structure of the different signal sequences. P_{snr} is independent of transmission structure. If the transmission bit rate B_r is known and the energy is spread over both the reference and the signal sequences then

$$\frac{E_b}{N_0} = \frac{P_{snr}}{2\tau B_r} \tag{48}$$

For orthogonal minimal constellations the bit rate, including the reference sequence time is given by

$$B_r = \frac{m}{2n\tau} \tag{49}$$

where τ is the sampling time of the system. Substituting equation 48 into 49 and rearranging gives

$$P_{snr} = \frac{m}{n}\cdot\left(\frac{E_b}{N_0}\right) \tag{50}$$

yielding

$$\hat{\mathbf{c}} = \overline{\mathbf{U}}^T\left(\left[\frac{\mathbf{I}_m}{\mathbf{0}_{n-m,m}}\right]\mathbf{c}+\frac{1}{\sqrt{m\left(\frac{E_b}{N_0}\right)}}\boldsymbol{\varepsilon}\right) \tag{51}$$

VII. SIMULATION CASE STUDY

Investigation into the "optimal" dimensionality of the new method [10] has shown that the optimal value for the scheme's dimension is approximately seven. This is for the given set of assumptions. For dimensions greater than seven, the improvement decreases, but the new scheme always performs better than any two dimensional quadrature scheme. The improvement decrease varies as a function of the relationship between the volume of the communication space and the surface area of the hypersphere, where the symbolic constellations lie. In addition, with higher dimensions, the computational complexity increases approximately as the third order of the dimension.

A case study for the proposed scheme is presented for a dimension of $m = 4$. This dimension has been chosen because it allows a clear demonstration of the advantages of using the scheme, whilst not presenting information that may be too confusing or complex for the purposes of illustration. The section is divided into two parts, section *A.* presents a transmission simulation where simulated real time random messages are transmitted and received with Gaussian White noise added in the communication channel. Illustrated, are the actual transmitted and received messages and the errors in the decoding of the information, due to the added noise. It can be demonstrated that, with this communication scheme, the error rates can be restored by an increase in the signal power. Section *B.* presents the 'Bit Error Rates' (BER) in terms of the 'Energy per Bit divided by the Noise Power' $\left(\frac{E_b}{N_0}\right)$. To find the BER for the estimator, the probability of the vector estimate lying outside its permitted region must be determined, that is that the following probability must be found

$$P(\hat{\mathbf{c}}_k \notin R^m(\mathbf{c}_k)) = P(E_\mu) \quad \forall \quad k \in [1,2^m] \tag{52}$$

For the purposes of the case study the chaotic system used is the Lorenz system. The advantages of this system are that it is simple, and has sufficiently chaotic behaviour for the purposes of demonstrating the communication scheme; but has characteristics that illustrate the problems that systems with a degree of periodicity can cause orthogonally oriented communication schemes. The system equations used for these results are

$$\begin{aligned} \alpha\dot{x} &= -\sigma x + \sigma y \\ \alpha\dot{y} &= rx - y - xz \\ \alpha\dot{z} &= xy + \beta z \end{aligned} \tag{53}$$

where $r = 28$, $\sigma = 10$ and $\beta = 8/3$. The constant α can be chosen to suit the sampling time that the particular system requires. For the following simulations $\alpha = 1$, all sampling times are assumed to be units of the chosen sample period and the first state $x(t)$ is used as the signal to be sampled.

A. Transmission Simulations

The next two figures 5 and 6, illustrate the results of simulating the communication scheme for a signal to noise ratio $P_{snr} = 10.0$. The simulation uses chaotic sequences that have not been enhanced by the matrix conditional method of sequences selection. Hence the noise rejection is improved only by an increase in the signal to noise ratio P_{snr}. Graph (a) shows a single vector sequence **x,** and graph (b) the persistent matrix **X** of zero mean sampled sequences, generated from the chaotic process. Graph (c) shows the orthonormal basis matrix sequences **U**, generated from the matrix **X** which in turn, when multiplied by the diagonal power balancing matrix **P**, give rise to the **Q** matrix which is used for encoding the signals sequence vector **s** shown in graph (e). In this scheme, the references are generated from the **x** vector by normalizing and power balancing it with the power value *p,* to generate the streamed and transmitted **z** vector. When these sequences are resampled at the receiver, they have been contaminated by noise, and are assembled into a persistent $\overline{\mathbf{Z}}$ matrix as shown in graph (d). The **Q** matrix is now encoded with a single symbol vector to generate the **s** vector of graph (e), and this is streamed and received in the same way as the **z** vector, to yield the noise contaminated $\bar{\mathbf{s}}$ vector in graph (f). The main advantage to this scheme is, that for each reference sequence and encoded sequence, there are *m* bits of information transmitted, which is one symbol representing *m* bits.

The same set of transmitted and received 'four' bit messages are shown in figures 6 graphs (a) and (b). Graph (c) demonstrates that there are no errors between the transmitted and the received message sequence for the chosen signal to noise ratio P_{snr} when the time delay is accounted for. The first *m* signals are always in error because the persistent matrices are not fully populated until four message sequences have been transmitted. This illustrates, as with all communication schemes, a need for some form of preamble before real information can be transmitted. The noise rejection can again be increased by the careful selection of the chaotic sequences. The limits of this scheme are the noise rejection due to the sequence length *n,* the chaotic sequence conditional selection and the dimension chosen for *m*, which is investigated and optimally selected in [10].

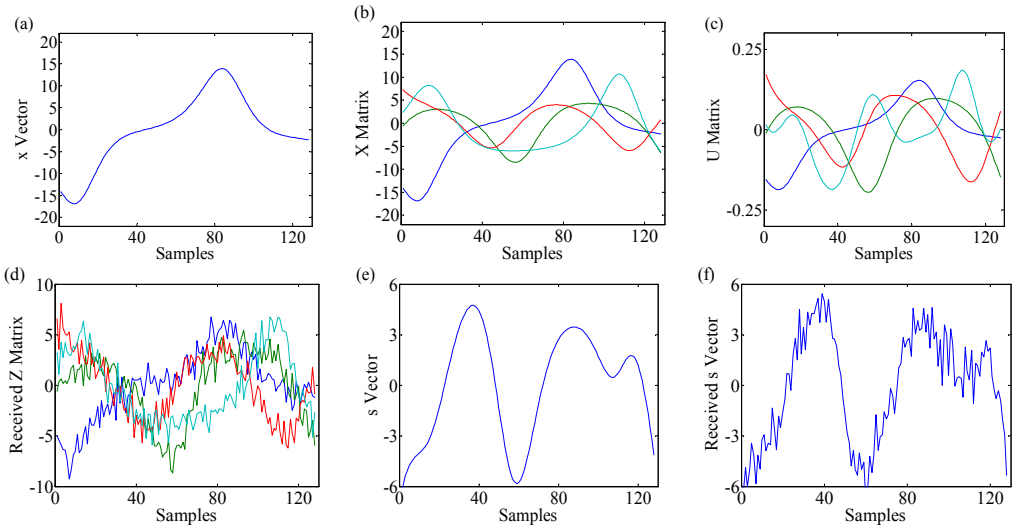

Figure 5: Indirect Persistent 'x' Scheme System Transmission Signals. $n = 128$, $m = 4$ and Power of Signal to Noise Ratio = 10.0

(a) Transmitter zero mean chaotic sequences **x**, (b) Persistent chaotic sequences **X**, (c) Generated orthogonal reference sequences **U**, (d) Received power balanced reference sequences $\overline{\mathbf{Z}}$, (e) Transmitted encoded signal sequence **s**, (f) Received encoded signal sequence $\overline{\mathbf{s}}$

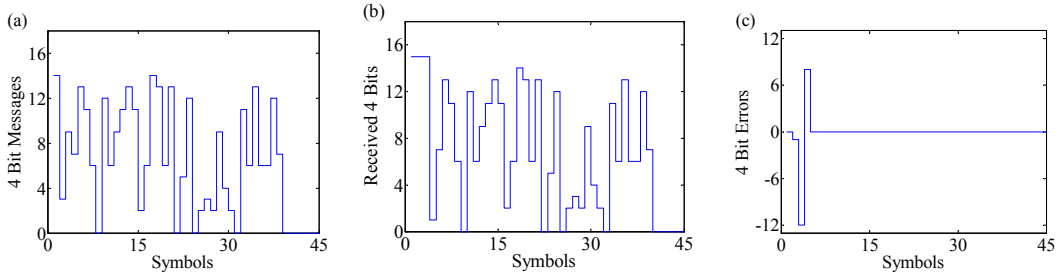

Figure 6: Indirect Persistent 'x' Scheme System Message Transmissions. $n = 128$, $m = 4$ and Power of Signal to Noise Ratio = 10.0, (a) Transmitted 4 bit message for encoding, (b) Received decoded 4 bit message, (c) Transmitted/Received 4 bit message delayed error

B. BER Simulations

For this scheme the **U** reference matrix is orthogonal. The DCSK BER of figure 7 (a) show better rates than the QCSK 16 example of graph (b). The OCVSK 16 example in graph (c) clearly outperforms the QCSK 16 scheme. The comparative graph in figure 8 shows that the Orthogonal Chaotic Vector Shift Keying (OCVSK) method have BER characteristics equivalent to Differential Chaos Shift Keying (DCSK), when the reference signals are orthogonal. The BER for the DCSK and the OCVSK 16 schemes have a dimensionally larger number of samples for the same BER, but the OCVSK 16 scheme has four times the data rate.

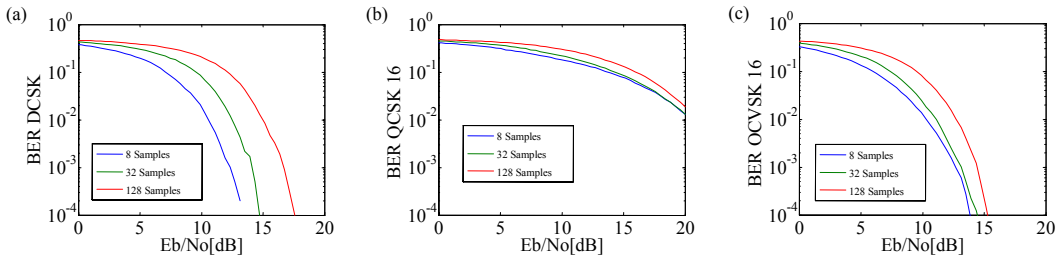

Figure 7: BER v $\frac{E_b}{N_0}$ Plots for direct 'm' Symbol 'U' Scheme for $n \in [8,32,128]$ samples

(a) DCSK, (b) QCSK 16 Symbol Constellation and (c) OCVSK 16

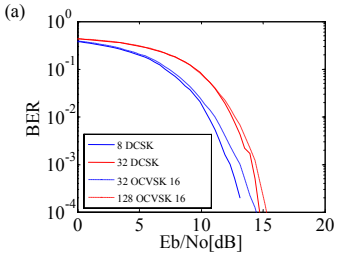

Figure 8: BER v $\frac{E_b}{N_0}$ Plots for direct '*m*' Symbol 'U' Scheme Comparison for DCSK and OCVSK 16 Showing that error rates are equivalent.

C. Non Orthogonal Case

As the signal references become more non-orthogonal, represented by non diagonal values of the signal characteristic matrix **W**, the BER graph diverges quite markedly. This is illustrated in figure 10. The generalized data rate for the orthogonal scheme is $\frac{m}{2n\tau}$ where *m* is the scheme dimension. Clearly as *m* approaches *n* the data rate tends towards the Shannon capacity [9] for this type of scheme. However, as *m* increases the BER steadily degenerates so the maximum channel capacity is not realizable. An optimum value for *m* has been conjectured in [10] for schemes of constant radius which has been found to be a value of $m = 7$.

The following case shows BER graphs for a specific characteristic **W** matrix. The **W** matrix essentially is a measure of how non orthogonal the **Z** matrix reference signal sequences is. All forms of the **W** matrix are upper triangular, which is a consequence of, the characteristic of the Gram-Schmidt orthonormalization process.

In this case the **W** matrix is given by

$$\mathbf{W} = \begin{bmatrix} 1 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{4}} \\ 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{4}} \\ 0 & 0 & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{4}} \\ 0 & 0 & 0 & \frac{1}{\sqrt{4}} \end{bmatrix} \qquad \mathbf{W}^T\mathbf{W} = \begin{bmatrix} 1 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} & \frac{1}{2} \\ \frac{1}{\sqrt{2}} & 1 & \frac{\sqrt{2}}{\sqrt{3}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{3}} & \frac{\sqrt{2}}{\sqrt{3}} & 1 & \frac{\sqrt{3}}{2} \\ \frac{1}{2} & \frac{1}{\sqrt{2}} & \frac{\sqrt{3}}{2} & 1 \end{bmatrix} \qquad (54), (55)$$

The non-orthogonal nature of this **W** matrix represents a set of signals, with a power of unity, each oriented at an angle of $\frac{\pi}{4}$ radians to its immediate predecessor. It represents a banded orthogonality shown by the $\mathbf{W}^T\mathbf{W}$ product of equation 55. This gives a greater span of BER shown in figure 9 (c) for signal to noise ratio P_{snr} over that of the orthogonal equivalent in figure 10, and illustrates that the banded non-orthogonality characteristic can be overcome by an increase in the number of samples *n*. The QCSK 16 examples of figure 9 (b) use only the first two columns and rows of the **W** matrix. Consequently, the effects of non-orthogonality become more apparent.

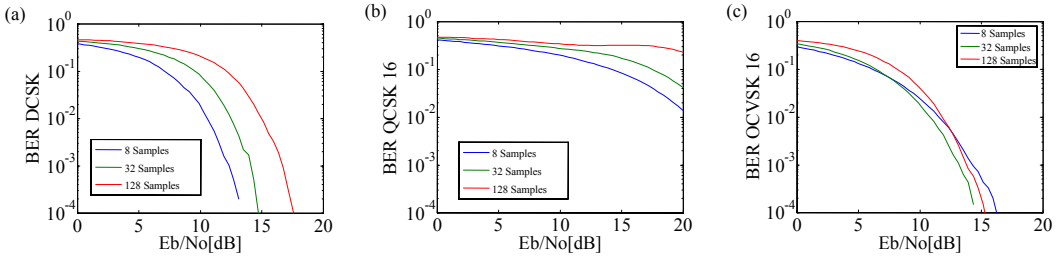

Figure 9: BER v $\frac{E_b}{N_0}$ Plots for W Scheme for $n \in [8,32,128]$ samples
(a) DCSK, (b) QCSK 16 Symbol Constellation and (c) OCVSK 16

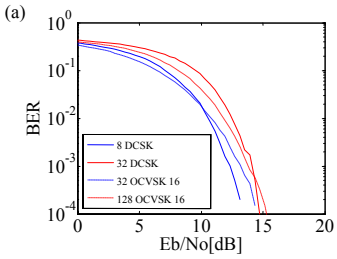

Figure 10: BER v $\frac{E_b}{N_0}$ $\frac{E_b}{N_0}$ Plots for W Scheme Comparison for DCSK and OCVSK 16 Scheme
Showing that the error rates are beginning to diverge.

VIII. CONCLUSIONS

The OCVSK method has demonstrated improvements in the robustness and security of communications links over those that are already presented in the literature. The structure of the scheme has improved the noise rejection because the effective "inter-symbolic distances" have been increased by considering a multi-dimensional paradigm rather than an increasing complex two dimensional one. The increase in dimensionality and the demonstration that an extension to the Fourier methods of QCSK is not applicable has required the solution to the problem of finding a method of producing multi-dimensional orthogonal signals. A novel method for this has been presented which is dependent on the vector sampling and subsequent orthonormalization of a set of vectors via the Gram-Schmidt process. A specific structural scheme called the "Indirect Persistent '**x**' Scheme" has been presented which has the merits of relative simplicity whilst maintaining the previously stated multi-dimensional benefits. The improvement in performance can be estimated by the ratio of inter-symbolic distances between two schemes. The inter-symbolic distance for all dimensions of the simple OCVSK structure is always $\sqrt{2}$ whereas, that of QCSK is symbol and hence dimension dependent. The performance increase factor is given by $\sqrt{\frac{1}{1-\cos\frac{2\pi}{m}}}$ where m is the dimension. For the presented scheme the improvement in performance is approximately 3.62 times.

The BER characteristics of the scheme have been shown to be equivalent to that of DCSK scheme and the data transmission rates are linearly dependent on the chosen dimension. A generic method of characterizing the noise contamination within schemes has be presented and successfully applied to the OCVSK scheme. Correct signal to noise characteristic results have been produced when this generic method of characterisation has been applied to three specific communications structures namely the OCVSK, QCSK and DCSK schemes.

APPENDICES

Appendix A: Quadrature Chaos Shift Keying Theory

Consider a signal on a closed interval $x(t)\ \forall\ t \in [0,T]$, which is generated by a chaotic system and is modified by removing the mean value so that it is a zero mean process, that is

$$\frac{1}{T}\int_0^T x(t)dt = 0 \qquad \text{(A1)}$$

then, if it admits to a Fourier expansion it can be expressed as

$$x(t) = \sum_{m=1}^{\infty} f_m \sin(m\omega t + \phi_m) \qquad \text{(A2)}$$

where $\omega = 2\pi / T$ and $f_0 = 0$.

Define the average power of this signal as

$$P_x = \frac{1}{T}\int_0^T x^2(t)dt \tag{A3}$$

which, because of the following properties of sinusoidal functions

$$\frac{1}{T}\int_0^T f_m \sin(m\omega t + \phi_m - \alpha) f_n \sin(n\omega t + \phi_n - \beta)dt$$

$$= \frac{1}{2} f_m^2 \cos(\alpha - \beta) \quad \forall \quad m = n$$

$$= 0 \qquad\qquad \forall \quad m \neq n \tag{A4}$$

can be expressed as

$$P_x = \frac{1}{2}\sum_{m=1}^{\infty} f_m^2 \tag{A5}$$

Now to derive a signal that is orthogonal to $x(t)$ apply a Hilbert Transform. This applies a phase shift of $\pi/2$ to every frequency in the signal. It can be achieved by taking a Fourier Transform of the signal and rotating the positive frequencies by $\pi/2$ and the negative ones by $-\pi/2$; finally inverting the resultant gives the transformed Fourier expansion

$$y(t) = \sum_{m=1}^{\infty} f_m \sin(m\omega t + \phi_m + \frac{\pi}{2}) \tag{A6}$$

then

$$x \perp y \Leftrightarrow \frac{1}{T}\int_0^T x(t)y(t)dt = 0 \tag{A7}$$

and it follows that

$$P_x = P_y \Leftrightarrow \frac{1}{T}\int_0^T x^2(t)dt = \frac{1}{T}\int_0^T y^2(t)dt \tag{A8}$$

Appendix B: Non-applicability of Fourier Expansion

Consider again the chaotic signal $x(t)$ defined on the closed interval $[0,T]$, which has had the mean value removed, and thus can be considered as a zero mean process over the interval.

Suppose that it will admit to a Fourier expansion of infinite length and that the conditions described in equations A2-A5 are applicable. Then for a signal $y(t)$ to be orthogonal to $x(t)$ over the interval then

$$\frac{1}{T}\int_0^T x(t)y(t)dt = 0 \tag{B1}$$

Now as $x(t)$ is given as an infinite Fourier Expansion, then there are an infinite number of signals orthogonal to it derived by applying phase shifts of $\pm\frac{\pi}{2}$ to each sinusoidal element of $x(t)$. So

$$y(t) = \sum_{k=1}^{\infty} f_k \sin(k\omega t + \phi_k \pm \frac{\pi}{2}) \tag{B2}$$

In the definition of integral *I* of equation A4 consider

$\alpha = 0$ and $\beta = \pm\frac{\pi}{2}$

then

$$I = \frac{1}{T}\int_0^T f_k \sin(k\omega t + \varphi_k) f_m \sin(m\omega t + \varphi_m \mp \frac{\pi}{2})dt$$

$$I = 0 \qquad \text{for} \qquad k \neq m$$

$$I = \frac{f_k^2}{2}\cos(\pm\frac{\pi}{2}) = 0 \qquad \text{for} \qquad k = m \tag{B3}$$

So $I = 0$ in all cases and therefore $y(t)$ is orthogonal to $x(t)$ in all cases.

Consider now an approximation to $x(t)$ derived from a limited sum of q sinusoidal elements, that is

$$x(t) = \sum_{k=1}^{q} f_k \sin(k\omega t + \phi_k) \tag{B4}$$

It follows that there are 2^q signals orthogonal to the $x(t)$ approximation expressed as follows

$$y_p(t) = \sum_{k=1}^{q} f_k \sin(k\omega t + \phi_k + F(p,k)\frac{\pi}{2}) \tag{B5}$$

where $p \in [1,2^q]$ and $F(p,k)$ is a notional function of p and k that takes the values ± 1 and varies as a Gray scale from $y_p(t) \to y_{p+1}(t)$ so that only one phase change takes place at each step in the entire expansion. Consider then the question: Are the $y_p(t)\ \forall\ p \in [1,2^q]$ signals generated from $x(t)$ mutually orthogonal. To show this is not the case consider any two derived signals $y_i(t)$ and $y_j(t)$ so

$$I = \frac{1}{T}\int_0^T y_i(t) y_j(t) dt \tag{B6}$$

$$I = \frac{1}{T}\int_0^T \sum_{k=1}^{q} f_k \sin(k\omega t + \varphi_k + F(i,k)\frac{\pi}{2}).f_k \sin(k\omega t + \varphi_k + F(j,k)\frac{\pi}{2}) dt \tag{B7}$$

Since all other terms vanish then

$$I = \sum_{k=1}^{q} \frac{1}{2} f_k^2 \cos\left(\left(F(i,k) - F(j,k)\right)\frac{\pi}{2}\right) \tag{B8}$$

$$I = \sum_{k=1}^{q} \frac{1}{2} f_k^2 G(k) \qquad \text{where } G(k) = \pm 1 \tag{B9}$$

It is theoretically possible to find a limited set of mutually orthogonal functions, but it is dependent on the values of each f_k and as each chaotically generated signal set is different; it is non-trivial problem to find a set $f_k\ \forall\ k \in [1,q]$ which drives equation B9 to zero. So, generally it is true, that all the $y_p(t)\ \forall\ p \in [1,2^q]$ signals are not mutually orthogonal.

Appendix C: Generation of Orthogonal Signal Sequences

The real vectors $\mathbf{u}_i$ can be obtained from real orthogonal functions $u_i(t)$ by the following process.
Let $\mathbf{u}_i = [u_{1i} \cdots u_{ni}]^T$ and $u_{ji} = u_i((j-1)\tau + t_i)\ \forall\ j \in [1,n]$, $i \in [1,m]$, t_i is the initial time at the point of sampling each vector and τ is the sampling period. These are arranged into a matrix $\mathbf{U} = [\mathbf{u}_1 \cdots \mathbf{u}_m]$ where $\mathbf{U} \in R^{n \times m}$. For practical consideration any real function $u_i(t)$ can generate a vector of signal values $\mathbf{u}_i$ by sampling. Conversely, the real signal functions $u_i(t)$ can be recreated in a digital to analogue converter (DAC). So the problem is reduced to generating a set of real valued orthogonal vectors $\mathbf{u}_i \in R^n$ and $i \in [1,m]$ in order to be able to generate a set of real time orthogonal functions $u_i(t)\ \forall\ i \in [1,m]$ given on an interval $t \in [0, n\tau]$.

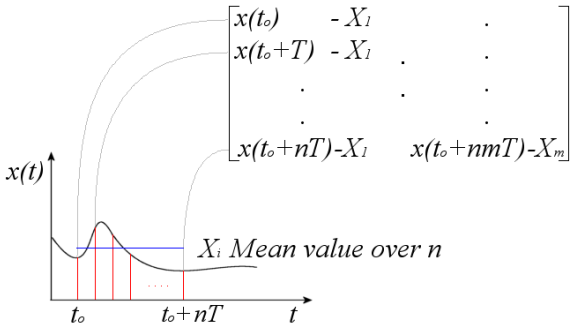

Figure C1: Signal Sampling to Matrix Concept

Consider a chaotic signal $x(t)$ sampled every τ seconds and the values placed into the columns of a matrix **X**. So $\mathbf{X} = [\mathbf{x}_1 \cdots \mathbf{x}_m]$ where $\mathbf{X} \in R^{nxm}$ and $\mathbf{x}_i \in R^n$ and in turn each $x_{ji} = x(((i-1)n + (j-1))\tau + t_0) \quad \forall \quad i \in [1,m],\ j \in [1,n]$ and t_0 is the initial sampling time. This is shown in figure C1. If the chaotic sequence is sufficiently varying then the $rank(\mathbf{X}) = m$ and the vectors that make up the matrix **X** will span an *m* dimensional subspace of the *n* space potentially spanned by a complete set of *n* vectors. A matrix **U** can be formed which is an orthonormal basis of **X** by using the Gram-Schmidt orthonormalization method. The result of the transformation is a simple linear transformation of **U** by an upper triangular matrix **W**. That is

$$\mathbf{X} = \mathbf{UW} \tag{C1}$$

where

$$\mathbf{U}^T\mathbf{U} = \mathbf{I}_m \tag{C2}$$

and **W** can be found as

$$\mathbf{U}^T\mathbf{UW} = \mathbf{U}^T\mathbf{X} \tag{C3}$$

So given equations C1-C2 it follows that

$$\mathbf{W} = \mathbf{U}^T\mathbf{X} \tag{C4}$$

Appendix D: Symbolic Encoded Vector Estimator

If the received signal is considered then

$$\bar{\mathbf{s}} = \mathbf{Qc} + \sigma\boldsymbol{\varepsilon} \tag{D1}$$

where $\boldsymbol{\varepsilon}$ is Gaussian White noise vector process with a zero mean and a unit variance, that is $\boldsymbol{\varepsilon}_i \sim N(1,0)\ \forall\, i \in [1,n]$, $E\{\boldsymbol{\varepsilon}\} = \mathbf{0}$ and $E\{\boldsymbol{\varepsilon}^T\boldsymbol{\varepsilon}\} = n$. In the following equations, the $\bar{}$ notation indicates a variable derived from received signal data and the ^ indicates an estimated value. So the signal estimate for a particular symbol represented by a received signal vector $\bar{\mathbf{s}}$ is given by

$$\hat{\mathbf{s}} = \bar{\mathbf{Q}}\hat{\mathbf{c}} \tag{D2}$$

expressing the error between the received signal vector and the estimated one as

$$\mathbf{e} = \bar{\mathbf{s}} - \hat{\mathbf{s}} \tag{D3}$$

and forming a squared error sum as

$$\eta = \mathbf{e}^T\mathbf{e} \tag{D4}$$

which can now be minimized with respect to the estimate of the encoding vector $\hat{\mathbf{c}}$ so

$$2\mathbf{e}^T\frac{\partial \mathbf{e}}{\partial \hat{\mathbf{c}}} = \mathbf{0}^T \tag{D5}$$

and from equations D2 and D3

$$\frac{\partial \mathbf{e}_i}{\partial \hat{\mathbf{c}}_i} = -\overline{\mathbf{Q}} \tag{D6}$$

Therefore equation D5 can be rearranged, incorporating equations D2 and D3 as

$$\overline{\mathbf{Q}}^T(\overline{\mathbf{s}} - \overline{\mathbf{Q}}\hat{\mathbf{c}}) = \mathbf{0} \tag{D7}$$

And finally forming an estimate of the encoding vector by rearranging equation D7 as

$$\hat{\mathbf{c}} = \left[\overline{\mathbf{Q}}^T\overline{\mathbf{Q}}\right]^{-1}\overline{\mathbf{Q}}^T\overline{\mathbf{s}}$$

Appendix E: Signal Characterization Derivation

A simple way of characterizing the effect of the Gram-Schmidt process is to consider how the transmitted matrix **Z** is constructed.

Firstly in order to complete this characterization we need to consider the diagonal power balancing matrix **P** in terms of the Signal Noise Ratio Power P_{snr}. Define the P_{snr} as

$$P_{snr} = \frac{\|\mathbf{z}\|^2}{\sigma^2} \tag{E1}$$

where $\|\mathbf{z}\|^2$ is the power of an arbitrary signal vector **z** and can be considered as the **2-**norm of the length *n* vector given as

$$\|\mathbf{z}\|^2 = \frac{1}{n}\mathbf{z}^T\mathbf{z} \tag{E2}$$

Now we require that

$$\overline{\mathbf{q}}^T\overline{\mathbf{q}} = n\|\mathbf{z}\|^2 \tag{E3}$$

that is, it is required that the transmitted vectors have a nominal power $\|\mathbf{z}\|^2$ and the total energy content increases linearly as the vector length *n.* For this scheme from equation 39

$$\overline{\mathbf{q}}_i^T\overline{\mathbf{q}}_i = p^2\overline{\mathbf{u}}_i^T\overline{\mathbf{u}}_i = p^2 \quad \forall \quad i \in [1, m] \tag{E4}$$

Combining equations E1, E3 and E4 yields

$$p = \sigma\sqrt{nP_{snr}} \tag{E5}$$

and therefore a characterized **P** can be formed as

$$\mathbf{P} = \sigma\sqrt{nP_{snr}} \cdot \mathbf{I}_m \tag{E6}$$

Now the **Z** matrix is a power balanced version of a normalized set of signal vectors generated by a chaotic process. As such they are not inherently orthogonal, and can be considered as the result of an upper-triangular linear transformation of an orthonormal basis, as described by equations 16 to 19 that is

$$\begin{aligned}\mathbf{Z} &= \mathbf{XP} \\ &= \sigma\sqrt{nP_{snr}} \cdot \mathbf{UW}\end{aligned} \tag{E7}$$

So finally the $\overline{\mathbf{Q}}$ matrix can be characterized in the following way if $\mathbf{G}(\overline{\mathbf{Z}})$ is the Gram-Schmidt matrix function and **W** is upper triangular with

$$\mathbf{w}_i^T\mathbf{w}_i = 1 \quad \forall \quad i \in [1, m] \tag{E8}$$

it follows that

$$\overline{\mathbf{U}} = \mathbf{G}\left(\sigma\left(\sqrt{nP_{snr}} \cdot \mathbf{UW} + \mathbf{E}\right)\right) \tag{E9}$$

Where **E** is a matrix of Gaussian White noise signals of unit variance with the same dimensions as **Q**. So finally

$$\overline{\mathbf{Q}} = \sigma\sqrt{nP_{snr}} \cdot \mathbf{G}\left(\sigma\left(\sqrt{nP_{snr}} \cdot \mathbf{UW} + \mathbf{E}\right)\right) \tag{E10}$$

the standard deviation constant σ within the Gram-Schmidt matrix function has no effect here, since it merely changes the component vector lengths and not their relationship to one another. Therefore this can be written as

$$\overline{\mathbf{Q}} = \sigma\sqrt{nP_{snr}} \cdot \mathbf{G}\left(\sqrt{nP_{snr}} \cdot \mathbf{UW} + \mathbf{E}\right) \tag{E11}$$

Likewise, the symbol signal vector can be characterized in the same, that is

$$\overline{\mathbf{s}} = \mathbf{Qc} + \sigma\boldsymbol{\varepsilon} \quad \forall \quad i \in [1, m] \tag{E12}$$

$\boldsymbol{\varepsilon}$ is an n length vector of Gaussian White noise signals of unit variance and so substituting equations 41 and E6 into this equation gives

$$\overline{\mathbf{s}} = \sigma\left(\sqrt{nP_{snr}} \cdot \begin{bmatrix} \mathbf{I}_m \\ \hline \mathbf{0}_{n-m,m} \end{bmatrix}\mathbf{c} + \boldsymbol{\varepsilon}\right) \tag{E13}$$

REFERENCES

[1] Williams C.: 'Robust chaotic communications exploiting waveform diversity. Part 1: Correlation detection and implicit coding', IET Communications, Nov. 2008, Vol. 2, Issue 10, pp. 1213-1222

[2] Williams C.: 'Robust chaotic communications exploiting waveform diversity. Part 2: Complexity reduction and equalisation', IET Communications, Nov. 2008, Vol., Issue 10, pp. 1223-1229

[3] Yu W., Morales A. and Fernandez G.: 'Robust chaotic communication via high gain observer', International Journal of Systems, Control and Communications, 2008, Vol. 1, Issue 2, Nov. 2008, pp. 179-192

[4] Chen, S. L., Chang S. M., Lin W. W. and Hwang T. T.: 'Digital Secure-Communication Using Robust Hyper-Chaotic Systems', International Journal of Bifurcation and Chaos, Nov. 2008, Vol. 18, Issue 11, pp. 3325-3339

[5] Alvarez J., Puebla H. and Cervantes I.: 'Stability of observer-based chaotic communications for a class of Lur'e systems', Int. J. Bifurcation and Chaos, 2002,Vol. 12, No. 7, pp.1605-1618

[6] Ricardo F., Ramón J. and Gualberto S.: 'A Chaos-Based Communication Scheme via Robust Asymptotic Feedback', IEEE Transactions on Circuits and Systems—I: Fundamental Theory and Applications, Oct. 2001, Vol. 48, No. 10, pp. 1161-1170

[7] Galias Z. and Maggio G. M.: 'Quadrature Chaos-Shift Keying', IEEE Transactions on Circuits and Systems I Fundamental Theory and Applications, 2001, Vol. 48, Issue 12, pp. 1510-1519

[8] Kolumbán G., Kennedy M. P. and Chua L. O.: 'The Role of Synchronization in Digital Communications Using Chaos – Part II: Chaotic Modulation and Chaotic Synchronization', IEEE Transactions on Circuits and Systems I Fundamental Theory and Applications, 1998, Vol. 45, Issue 11, pp. 1129-1140

[9] Dunlop J. and Smith D.: 'Telecommunications Engineering', (Van Nostrand Reinhold (UK) Co. Ltd., 1984)

[10] Wren T. J.: 'Orthogonal Chaotic Vector Shift Keying in Digital Communications', DPhil Thesis, 2007

[11] Peebles Jr. P. Z.: 'Digital Communications Systems', (Prentice Hall International Inc., 1987)

[12] Schwartz M., 'Information Transmission Modulation and Noise', (McGraw-Hill, 1980, 3rd Edn.)

[13] Pecora L. M. and Carroll T. L.: 'Synchronization in Chaotic Systems', *Phys. Rev. Lett.*, 1990, Vol. 64, pp. 821-824

[14] Pecora L. M. and Carroll T. L.: 'Driving Systems with Chaotic Signals', *Phys. Rev. A*, 1991, Vol. 44, pp. 2374-2384

[15] Cuomo K. M. and Oppenheim A. V.: 'Circuit Implementation of Synchronized Chaos with Applications to Communications', 1993, *Phys. Rev. Lett.*, Vol. 71, pp. 65-68

[16] Pecora L. M. and Carroll T. L.: 'Synchronizing Hyperchaotic Volume-Preserving Maps and Circuits', IEEE Transactions on Circuits and Systems I Fundamental Theory and Applications, 1998, Vol. 45, Issue 6, pp. 656–659

[17] Kolumbán G., Kennedy M. P., and Chua L. O.: 'The role of synchronization in digital communication using chaos—Part I: Fundamentals of digital communication', *IEEE Trans. Circuits Syst. I*, Oct. 1997, vol. 44, pp. 927–935

[18] Kolumbán G., Kennedy M. P., and Chua L. O.: 'The role of synchronization in digital communication using chaos—Part II: Chaotic modulation and chaotic synchronization', *IEEE Trans. Circuits Syst. I*, Nov. 1998, Vol. 45, pp. 1129–1140

[19] Kolumbán G., Kis G., Jákó Z., and Kennedy M. P.: 'FM-DCSK: A robust modulation scheme for chaotic communications', *IEICE Trans. Fund. Electron., Commun. Comput. Sci.*, 1998, Vol. E-81A, No. 9, pp.1798–1802

[20] Couch L.W.: '*Digital and Analog Communication Systems',* (Prentice-Hall, 1997)

[21] Kolumbán G., M. P. Kennedy, and Kis G.: 'Multilevel Differential Chaos Shift Keying', in *Proc. Int. Workshop, Nonlinear Dynamics of Electronics System NDES'97*, 1997, pp. 191–196

[22] Kennedy M. P., Kolumbán G., Kis G., and Jákó A.: 'Recent advances in communication with chaos', in *Proc. IEEE Int. Symp. on Circuits and System ISCAS'98*, 1998, pp. 461–464

[23] Jákó Z.: 'Performance improvement of DCSK modulation', in *Proc. Int. Workshop, Nonlinear Dynamics of Electronics Systems NDES'98*, 1998, pp. 119–122

[24] Lee E. A. and Messerschmitt D. G.: '*Digital Communications'*, (Kluwer Academic, 1993, 2nd Edn.)